AD

TECHNICAL REPORT ARBRL-TR-02317

PARTITIONABILITY OF IMPLICIT LEAST
SQUARES MODEL FITTING PROBLEMS

Aivars Celmiņš

May 1981

US ARMY ARMAMENT RESEARCH AND DEVELOPMENT COMMAND
BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND

Approved for public release; distribution unlimited.

81 6 17 008

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TECHNICAL REPORT ARBRL-TR-02317	2. GOVT ACCESSION NO. AD-A101 356	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Partitionability of Implicit Least Squares Model Fitting Problems		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Aivars Celmiņš		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS US Army Ballistic Researh Laboratory ATTN: DRDAR-BLI Aberdeen Proving Ground, MD 21005		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 1L161102AH43
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Armament Research and Development Command US Army Ballistic Research Laboratory ATTN: DRDAR-BL Aberdeen Proving Ground, MD 21005		12. REPORT DATE MAY 1981
		13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) (dll)

Non-linear least squares model fitting problems differ from general optimization problems due to the special structure of the least squares normal equations. In simple least squares problems, that structure allows one to partition the normal equation system, whereby significant algorithmic simplification can be achieved. More general least squares problems can often be made partitionable by proper parameter manipulations. In this report, sufficient partitionability conditions are derived and parameter manipulations are discussed for the achievement of partitionability.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

TABLE OF CONTENTS

Accession For

NTIS GRA&I	✔	
DTIC TAB	☐	
Unannounced	☐	
Justification		

By

Distribution/

Availability Codes

Dist	Avail and/or Special
A	

3/4

1. INTRODUCTION

We consider least squares model fitting tasks that can be formulated as constrained minimization problems. In general, the constraints, i.e., the model equations, are implicit nonlinear relations between observables and parameters. The solution of such problems can be found by the Lagrange multiplier technique which provides a system of coupled nonlinear normal equations. The equations can be used to analyze the sensitivity of the solution to data perturbations, and to obtain numerical values of optimal residuals and parameters. However, the numerical solution of the system is not necessarily trivial, because the size of the system is proportional to the number of data. Because of the potentially very large size of the equation system, a partitioning, if possible, has many algorithmic advantages. Fortunately, many typical least squares model fitting problems have such a norm. equation structure that a partitioning of the equation system can be achieved by proper manipulations of the model equations. Particularly effective can be a manipulation of parameters, e.g., an introduction of new parameters and/or a formal elimination of some parameters. We will consider the effects of such manipulations and derive partitionability conditions which can be used as a basis for a rational choice of model equation formulation, and for a rational planning of experiments.

In Section 2 we will give a formal definition of the least squares model fitting problem, and establish the normal equations. Algorithms for the numerical solution of the equations will be outlined in Section 3, where also the sensitivity of the solution to data perturbations is investigated. Partitionability of the normal equations and corresponding parameter manipulation are discussed in Sections 4 and 5, repectively. Section 6 brings an example for partitioning of normal equations arising in the adjustment of a planimetric traverse.

2. THE MODEL FITTING PROBLEM

Let the mathematical model of an observable event be formulated by a set of r independent equations, say,

$$F(x,t) = 0 , \tag{2.1}$$

where $F \in R^r$ is a vector function, $x \in R^n$ represents potential observations, and $t \in R^p$ is a vector of free parameters. Eqs. (2.1) may be prescribed by a theory of the event, or chosen by other considerations. Regardless of their source, the equations are considered as a given description of the event. A model fitting problem arises when one seeks to determine the validity of the mathematical description by testing Eqs. (2.1) with experimental data. In such tests, typical magnitudes of the dimensions p, r, and n are 10, 100, and 1000, respectively. We assume that the dimensions satisfy the inequalities

5

$$0 \leq p < r \leq n, \tag{2.2}$$

which assure that the optimization problem has sufficient degrees of freedom.

We restrict our considerations to model functions F that are twice differentiable with respect to all its $n + p$ arguments. Differentiability of model functions is typical for many problems in engineering, physics and other fields of application. For the present analysis, it has to be assumed only within a neighborhood of the adjusted observations x and of the optimal parameter t.

Let $X \in R^n$ be the vector of the actual observations. Because X contains observational inaccuracies one cannot expect that the theoretical description of the event, i.e., Eq. (2.1), is satisfied at $x = X$. Instead, one needs corrections (residuals) c which are added to the observations so that Eq. (2.1) is replaced by

$$F(X + c, t) = 0 \quad . \tag{2.3}$$

Eq. (2.3) means that we expect the theoretical relations between observables and parameters to be satisfied at a vicinity $X + c$ of the actual observations X.

The mathematical goal of model fitting is to find residuals c and parameters t that are optimal in some sense. We do not put any restrictions on t, but obviously would like the residuals c to be as small as possible. Hence, a general model fitting problem can be formulated as the following constrained minimization task

$$||c|| = min. \quad , \tag{2.4a}$$

$$F(X + c, t) = 0 \quad . \tag{2.4b}$$

The solution of this problem, of course, depends on the definition of the norm $||c||$. In this article we consider only elliptic vector norms, defined by

$$||c|| = \sqrt{c^T M c} \quad , \tag{2.5}$$

6

where M is a positive definite norm matrix. One reasonable choice of M is a diagonal matrix where the elements are proportional to the inverse squares of estimated standard errors of the observed components of X.[1,2] With this choice, the norm $||c||$ becomes dimensionless and the corresponding minimization problem is called "weighted least squares." A more general choice for M is the inverse of an estimated variance-covariance matrix R of the observations X.[3,4] If the observational errors of X are normally distributed, then the use of this norm, i.e., of

$$||c|| = \sqrt{c^T R^{-1} c} \tag{2.6}$$

in Eq. (2.4a) produces a maximum likelihood result.[5] The norm (2.6) is, of course, also dimensionless, and it includes the weighted least squares norm as a special case.

Using the norm (2.6) we formulate a general least squares model fitting problem as follows:

$$W = ||c||^2 = c^T R^{-1} c = \min. , \tag{2.7a}$$

$$F(X + c, t) = 0 . \tag{2.7b}$$

The unknowns of the problem are the residuals c and the parameter vector t. Given are the observations X, the model function F and an estimated variance-covariance matrix R. The latter need to be known only

---

[1] W.E. Deming, "Statistical Adjustment of Data," John Wiley & Sons, New York, NY, 1944.

[2] S. Brandt, "Statistical and Computational Methods in Data Analysis," North-Holland Publishing Co., Amsterdam, 1970.

[3] D. Brown, "A Matrix Treatment of Least Squares Considering Correlated Observations," USA Ballistic Research Laboratories Report No. 937, 1955. (AD #71209)

[4] J.M. Tienstra, "Theory of the Adjustment of Normally Distributed Observations," N.V. Uitgevery "Argus," Amsterdam, 1956.

[5] M.F. Britt and R.H. Luecke, "The Estimation of Parameters in Nonlinear, Implicit Models," Technometrics, 15, 233-247, 1973.

up to a factor, because the inclusion of an arbitrary factor in Eqs. (2.7a) does not change the minimum condition. Also, the model function F can be manipulated as long as the result produces a system of equations mathematically equivalent to Eqs. (2.7b).

Problem (2.7) can be reduced to an unconstrained minimization problem if the residuals c can be eliminated from Eqs. (2.7a) by using Eqs. (2.7b). This is the case, e.g., when Eqs. (2.7b) are explicit in terms of X + c. Many least squares algorithms have been devised to treat this special problem. In this article, we consider the more general situation where F is a general nonlinear function such that a formal elimination of all or some residuals is either not possible or not practical.

In order to simplify our notation in the subsequent analysis, we shall denote derivatives of F by subscripts. Thus, e.g.,

$$\frac{\partial F(X + c, t)}{\partial X} = F_X(X + c, t)$$

is a rxn matrix, and

$$\frac{\partial^2 (K^T F(X + c, t))}{\partial X \partial t} = (K^T F(X + c, t))_{Xt},$$

where $K \in R^r$, is a nxp matrix.

In addition to the differentiability of F we also assume that in a neighborhood of the solution $(X + c, t)$ the following rank conditions are satisfied:

$$\text{rank } F_X = r \tag{2.8}$$

and

$$\text{rank } F_t = p \ . \tag{2.9}$$

The condition (2.8) insures that the model Eqs. (2.3) are independent. The conditon (2.9) excludes model formulations with redundant parameters.

8

Next, we obtain normal equations for the problem (2.7) using Lagrange multipliers. Let $k \in R^r$ be a vector of correlates (Lagrange multipliers), and let

$$\hat{W} = 1/2\ c^T R^{-1} c - k^T F\ (X + c,\ t) \tag{2.10}$$

be the modified object function. By setting the derivatives of $\hat{W}$ with respect to c, t and k equal to zero, we obtain the equations.[5,6,7,8,9]

$$c - R \cdot F_X^T\ (X + c,\ t) \cdot k = 0\ , \tag{2.11a}$$

$$k^T \cdot F_t\ (X + c,\ t) = 0\ , \tag{2.11b}$$

$$F(X + c,\ t) = 0\ . \tag{2.11c}$$

The normal Eqs. (2.11) generally have more than one solution, and the solution of the optimization problem (2.7) is one of the several solutions of Eqs. (2.11). The selection of the proper solution is normally done by an analysis of problem related background information. We shall not discuss such analyses in this article, and concentrate instead on the finding of any numerical solution of Eq. (2.11).

3. ITERATION ALGORITHMS AND EFFECTS OF DATA PERTURBATION

The normal Eqs. (2.11) are nonlinear with respect to the unknowns c and t. Therefore, their numerical solution will generally require an iteration. One obtains second order Newton-type iteration equations by expanding the normal equations at an approximate solution.

---

[6] A. Celmiņš, "Least Squares Adjustment with Finite Residuals for Non-Linear Constraints and Partially Correlated Data," USA Ballistic Research Laboratories Report No. R-1658, 1973. (AD #766283)

[7] A.F. Pope, "Two Approaches to Nonlinear Least Squares Adjustments," The Canadian Surveyor, 28, 663-669, 1974.

[8] R.M. Passi, "Use of Nonlinear Least Squares in Meteorological Applications," Journal of Applied Meteorology, 16, 828-832, 1977, and 17, 1579-1580, 1978.

[9] W.H. Jefferys, "On the Method of Least Squares," The Astronomical Journal, 85, 177-182, 1980.

Let C, K and T be approximations to the solution vectors c, k and
t, respectively, and ε, κ and τ be the corresponding corrections. The
linear terms of an expansion of Eq. (2.11) at the approximate solution
produce the following system of Newton-Raphson equations for the cor-
rections[7]:

$$\left[I - R \cdot (K^T \cdot F)_{XX}\right] \cdot \varepsilon - R \cdot F_X^T \cdot (K+\kappa) - R \cdot (K^T \cdot F)_{Xt} \cdot \tau = -C$$

$$(K^T \cdot F)_{tX} \cdot \varepsilon + F_t^T \cdot (K+\kappa) + (K^T \cdot F)_{tt} \cdot \tau = 0 \qquad (3.1)$$

$$F_X \cdot \varepsilon + F_t \cdot \tau = -F$$

The arguments of F and of its derivatives in Eq. (3.1) are the approxi-
mations X + C and T.

An iteration based on Eq. (3.1) proceeds by computing the corrections
ε, κ and τ, adding them to the approximations C, K and T, respectively,
and repeating the process. The equations may be rearranged into a more
convenient form for the iteration. An example of such rearranged iter-
ation equations is given in the Appendix and corresponding computer
programs are described in Reference 10.

An often used variation of the Newton-Raphson equations is obtained
by setting in Eq. (3.1) all second order derivatives of F equal to
zero.[1,2,5,9] The resulting equations are called Gauss-Newton equations.
Iteration algorithms based on Gauss-Newton equations converge only
linearly and may have other disadvantages.[7]

The linear terms of the expansion of the normal equations, i.e.,
the equation system (3.1), also provide a means to obtain estimates of
the effects of data perturbations on the solution. To this end, we
express the normal equations in terms of the corrected observations
x = X + c, obtaining the system

$$x - R \cdot F_X^T(x,t) \cdot k = X ,$$

$$k^T \cdot F_t(x,t) = 0 , \qquad (3.2)$$

$$F(x,t) = 0 ,$$

[10] A. Celminš, "A Manual for General Least Squares Model Fitting,"
USA Ballistic Research Laboratory Report ARBRL-TR-02167, 1979.
(AD #B040229L)

which we expand at the solution. The linear terms of the expansion yield the following relation between the differentials of the solution x, k, t and the differentials of the observations X:

$$\left[I - R \cdot (k^T \cdot F)_{xx} \right] dx - R \cdot F_x^T dk - R \cdot (k^T \cdot F)_{xt} dt = dX \ ,$$

$$(k^T \cdot F)_{tx} dx + F_t^T dk + (k^T \cdot F)_{tt} dt = 0 \ , \qquad\qquad (3.3)$$

$$F_x dx + F_t dt = 0 \ .$$

The coefficient matrices in Eq. (3.3) are identical to those in Eq. (3.1), except that now the functions are evaluated at the solution. Therefore, differential changes of the solution corresponding to data perturbations dX can be calculated conveniently by using the iteration equations of the Appendix. Thus, if one is interested in the changes dt of the parameters corresponding to the perturbations dX, one can use the formula

$$N \ dt = S \ dX \ , \qquad\qquad (3.4)$$

where N and S are defined in the Appendix in terms of F and its derivatives.

Eq. (3.4) also can be used to derive a formula for an estimate of the variance-covariance matrix V_t of the components of the parameter vector t. The formula is obtained by applying the law of variance propagation to Eq. (3.4) with the result[6,7,10]

$$V_t = N^{-1} S R S^T (N^{-1})^T \ . \qquad\qquad (3.5)$$

In case the variance-covariance matrix R of the data has been estimated only up to a factor, the formula must be supplemented by an estimate of that factor. The usual estimate is

$$m_o^2 = \frac{1}{n-p} c^T R^{-1} c \ . \qquad\qquad (3.6)$$

The square root m_o of the factor is also called the standard error of weight one.

11

It is clear from the derivation of Eq. (3.5) that the formula for the variances of t contains first and second order derivatives of the model function F in spite of the fact that the formula is only a first order estimate of the variances. (The dependence on the second order derivatives is shown explicitly in the Appendix.) As noted in Reference 7, the reason for the presence of second order derivatives in Eq. (3.5) is the appearance of first order derivatives in the normal Eq. (3.2) which are perturbed and expanded when Eq. (3.5) is derived. Authors who prefer Gauss-Newton algorithms for the numerical solution of the normal equations tend to overlook this fact and present variance estimate formulas without second order derivatives of the model function F[1,2,5,8,9]. Such formulas are less than first order accurate and should not be used without an estimate of the effect of the neglected terms. One can easily construct examples where the second order derivative terms contribute significantly to V_t either increasing or decreasing the estimated variances.

4. PARTITIONABILITY

If the data volume is large, then the numerical solution of the normal Eqs. (2.11) can be a formidable task. In a typical model fitting problem the dimension n of the observations is the order 1000 or larger, and one has to manipulate matrices of the order n×n in the Newton-Raphson iteration equations. Fortunately, many least squares problems have such a structure that the large systems of equations can be partitioned into a set of smaller systems, whereby the manipulation of the large matrices can be avoided.

We illustrate this partitionability with a curve fitting problem in the y,z-plane. Let the curve be defined by the implicit equation

$$f(x;t) = f(y,z;t) = 0 , \qquad (4.1)$$

where x is the coordinate vector in the y,z-plane and t is a parameter vector. Let the observations X_i, $(i = 1,2,\dots,s)$ be the coordinates of s points in the y,z-plane, i.e.,

$$X_i = \begin{pmatrix} y_i \\ z_i \end{pmatrix} , \quad i = 1,2,\dots,s , \qquad (4.2)$$

and let the accuracies of the observed points be characterized by corresponding estimated variance-covariance matrices

$$R_i = \begin{pmatrix} v_{iyy} & v_{iyz} \\ v_{iyz} & v_{izz} \end{pmatrix} , \quad i = 1,2,\dots,s . \qquad (4.3)$$

12

The correspondences between this problem and the general model fitting problem are as follows

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_s \end{pmatrix} , \tag{4.4}$$

$$R = \begin{pmatrix} R_1 & & \\ & \ddots & 0 \\ 0 & & R_s \end{pmatrix} , \tag{4.5}$$

and

$$F(X,t) = \begin{pmatrix} f(X_1,t) \\ \vdots \\ f(X_s,t) \end{pmatrix} . \tag{4.6}$$

Using these correspondences, the curve or model fitting problem (2.7) can be defined in terms of the subsets X_i of X, the submatrices R_i of R, and the components f of F as follows:

$$W = \sum_{i=1}^{s} c_i^T R_i^{-1} c_i = \min.$$

$$f(X_i + c_i, t) = 0, \quad i=1,2,\ldots,s . \tag{4.7}$$

The normal equations of the optimization problem (4.7) are

$$c_i - R_i \cdot f_X^T (X_i + c_i, t) \cdot k_i = 0, \quad i=1,2,\ldots,s \tag{4.8a}$$

$$\sum_{i=1}^{s} k_i \cdot f_t (X_i + c_i, t) = 0 \tag{4.8b}$$

$$f(X_i + c_i, t) = 0, \quad i=1,2,\ldots,s \tag{4.8c}$$

where k_i $(i=1,2,\ldots,s)$ are the correlates of the problem.

13

A comparison of Eqs. (4.8) with the general normal Eqs. (2.11) shows that in the curve fitting problem the large equation system is partitioned into a set of smaller systems, so that the largest dimension of matrices to be manipulated is the maximum of 2x2 and pxp. Particularly, Eq. (4.8a) is a set of s systems of two equations, each system depending only on two residuals, whereas, Eq. (2.11a) is one system of 2s equations depending on 2s residuals. Likewise, Eq. (4.8c) are s scalar equations, each depending on two distinct residuals, whereas the corresponding Eq. (2.11c) is a system of s coupled equations for all 2s residuals. Obviously, the numerical treatment of Eqs. (4.8) is much simpler than that of Eqs. (2.11).

A basic property of the sample problem (4.7) is that the r constraints are scalar equations, (hence $r_i = 1$ and $r = \Sigma r_i = s$) each depending on a distinct subset X_i of the observation vector X, and that the s subsets X_i are not correlated. We call such a problem a _standard_ _least squares problem_ because of its common occurance and simplicity. Standard least squares problems are easier to solve numerically than general problems, because the maximum dimensions of matrices in the normal equations are independent of the total number of observations. A problem with the latter property we call _totally partitionable_. Hence, a standard least squares problem is totally partitionable. If the data are not correlated, then any fitting of a hypersurface to points in a space of observables is a totally partitionable problem. Such a fitting in a, say, m-dimensional space is also a standard problem if the dimension of the hypersurface is m-1.

Next, we derive conditions for partitionability of the normal equations by comparing the structures of Eqs. (2.11) and (4.8). First, we notice that in order to be able to partition Eq. (2.11c) at all, the model function F must be transformable into such a form that subsets of components of F depend on distinct subsets of the observations X. This property can be conveniently expressed by the requirement that the Jacobian matrix $\partial F/\partial X$ has a stretched block diagonal structure, i.e.,

$$\frac{\partial F}{\partial X} = \begin{pmatrix} \dfrac{\partial F_1}{\partial X_1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{\partial F_s}{\partial X_s} \end{pmatrix} \qquad (4.9)$$

In the sample curve fitting problem (4.7), F has this property, whereby the submatrices $\partial F_i/\partial X_i$ are the two-component vectors $\partial f/\partial X_i$. In more general situations the submatrices have dimensions $r_i \times n_i$ and the r_i and n_i can be different for different indexes i. Because $\partial F/\partial X$ is a rxn matrix, then obviously $\Sigma r_i = r$ and $\Sigma n_i = n$.

14

The stretched block diagonal structure (4.9) of $\partial F/\partial X$ suffices to partition Eq. (2.11c). In order to partition Eq. (2.11a) too, one needs an additional condition on the variance-covariance matrix R. If R is diagonal and (4.9) holds, then Eq. (2.11a) is partitionable. However, for partitionability it is already sufficient if R has a block diagonal structure

$$R = \begin{pmatrix} R_1 & & 0 \\ & \ddots & \\ 0 & & R_s \end{pmatrix} \quad , \tag{4.10}$$

where the dimensions n_i of the submatrices R_i match the dimensions n_i of the submatrices $\partial F_i/\partial X_i$. Both of these conditions together are sufficient to partition the problem into s parts. Thus, if R and $\partial F/\partial X$ have the indicated structures, Eq. (2.11a) has the form

$$\begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} - \begin{pmatrix} R_1\left(\dfrac{\partial F_1}{\partial X_1}\right)^T & & 0 \\ & \ddots & \\ 0 & & R_s\left(\dfrac{\partial F_s}{\partial X_s}\right)^T \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_s \end{pmatrix} = 0 \quad , \tag{4.11}$$

where the c_i are distinct subsets of c with the dimensions n_i, and the k_i are correlate vectors with the dimensions r_i.

In summary, sufficient for the partitionability of a least squares model fitting problem is that the following two conditions hold:

a. R has a block diagonal structure (4.10), and

b. $\partial F/\partial X$ has a matching stretched block diagonal structure (4.9).

In data reduction problems one has no control over the structure of R, except during the planning stage of an experiment. Once the measurements are made, R is part of the given data basis. However, in many practical problems R is diagonal or nearly diagonal with few non-zero off-diagonal elements. In these cases, the partitionability of the problem depends on the formulation of the model equations F = 0. By a proper manipulation of the constraints one can often partition a problem that was not partitionable in the original formulation. We shall give an example of such a manipulation in Section 6.

15

The algorithmic advantages of partitioning cannot be overemphasized. In fact, partitionability rather than the special form of the object function W is the practically important difference between a least squares problem and a general optimization problem. Most published algorithms for the solution of least squares problems are restricted to partitionable cases.

5. MANIPULATION OF PARAMETERS

This section gives an overview of parameter manipulations that can be used to achieve partitionability of least squares model fitting problems. First, we express the constraint Eqs. (2.3) in the more convenient form

$$\overline{F} \ (c,t) \ = \ 0 \tag{5.1}$$

by including the observed X into the definition of the model function $\overline{F}$. Like Eq. (2.3), Eq. (5.1) is a set of r scalar equations. In general, each of the r equations depends on different subsets of the components of c and t. A subset t_e of the parameter vector t we define as the vector of _essential parameters_ if all components of t_e appear in all r Eqs. (5.1). All other parameters we call _non-essential_. A constraint formulation that contains only essential parameters we call a _minimal formulation_.

Minimal formulations of constraints are important in the context of partitionability. To illustrate this, we notice that the numbers and types of parameters are not intrinsic properties of a model fitting problem. Parameters can be eliminated and added, within limits, without changing the solution of the problem. However, from an algorithmic viewpoint it is not advisable to eliminate essential parameters, i.e., to reduce the problem formulation below a minimal formulation.

We illustrate this remark by considering a standard problem with a diagonal variance-covariance matrix R and only essential parameters. Let the constraints (5.1) of the problem be, componentwise,

$$\overline{f}_i \ (c_i,t) \ = \ 0, \quad i=1,2,\ldots,r \ . \tag{5.2}$$

In order to eliminate the parameters we may use the last p Eqs. (5.2) and express t in terms of the residual subset $c_e = (c_{r-p+1}, \ldots, c_r)$. Substituting this expression into the first r-p equations, one obtains a system of constraints, equivalent to the original system, but without parameters, namely,

16

$$\overset{\gamma}{F}(c) = 0 \tag{5.3}$$

with the components

$$\overset{\gamma}{f}_i(c_i, c_e) = 0, \quad i=1,\ldots,r-p , \tag{5.4}$$

Now, the arguments of the components $\overset{\gamma}{f}_i$ of $\overset{\gamma}{F}$ are not distinct subsets of c and, therefore, the Jacobian matrix $\partial \overset{\gamma}{F}/\partial c$ has not the necessary stretched block diagonal form. The problem is not partitionable in this formulation.

While a problem formulation with fewer constraints than in a minimal formulation certainly is not optimal for a numerical treatment, one may find, in some cases, that larger than minimal formulations are more practical. One obtains such formulations by introducing new parameters into the problem. The total number of parameters that can be added to a least squares problem is, however, limited by the inequalities (2.2). Let us assume that for a given problem the inequalities are satisfied, and let us introduce $\hat{p}$ new parameters into the problem. The corresponding $\hat{p}$ new equations which define the parameters are added to the set of constraint equations. Therefore, the inequality (2.2) for the new problem formulation is

$$0 \le p + \hat{p} < r + \hat{p} \le n . \tag{5.5}$$

Hence, the number $\hat{p}$ of new parameters that can be introduced into a problem is limited by the condition

$$\hat{p} \le n - r . \tag{5.6}$$

A "natural," application oriented, formulation of a model fitting problem is not necessarily the most advantageous one for numerical treatment, and may even be sub-minimal, as in the following example. Suppose that some effect y of an explosion has been observed at different stations as a function of time t. Then the data basis consists of an observation T_0 of the time t_0 of the explosion and of a series of pairs (T,Y), providing the observed effects Y at times T. Let the theoretical model equation of the observed effect at station j be

$$y = f_j(t - t_0; \alpha, \beta, \gamma) \tag{5.7}$$

17

where α, β, and γ are essential model parameters. Then the corresponding constraint equations are

$$Y_i + c_{Yi} - f_i(T_i + c_{Ti} - (T_o + c_{To}); \alpha, \beta, \gamma) = 0, \quad (i = 1, 2, \ldots, r) .$$

(5.8)

The formulation (5.8) is minimal, because all parameters are essential. However, the problem is not partitionable even when all observations are uncorrelated, because all constraint equations contain the same observation T_o.

A partitioning of this problem can be easily achieved by introducing the starting time of the explosion as a fourth parameter δ. The correspondingly modified set of constraint Eqs. (5.8) is

$$\left. \begin{array}{l} Y_i + c_{Yi} - f(T_i + c_{Ti} - \delta; \sigma, \beta, \gamma) = 0 , \quad (i = 1, \ldots, r), \\[2em] T_o + c_{To} - \delta = 0 . \end{array} \right\} (5.9)$$

Now we have $r + 1$ constraints, four parameters and $(m + 1) \cdot r + 1$ observations, where $m = \dim Y$. The inequalities (5.5) are in this case

$$0 \leq 4 < r + 1 \leq (m + 1) \cdot r + 1 , \tag{5.10}$$

indicating, that one needs at least four observation sets (T, Y) to have a regular adjustment problem.

The problem formulation (5.9) is not minimal, because the parameters α, β, and γ are not essential. However, because they appear in all but one of the constraint equations, their elimination would not simplify the problem. If the data are not correlated, then in the formulation (5.9) the problem is totally partitionable. It is, in fact, a standard problem, if the observed effects Y_i are scalar, i.e., if $m = 1$.

The goal of manipulation of the model equations is to obtain an equation system with a stretched block diagonal Jacobian matrix. If the model equations are linear, then this can be achieved by algebraic manipulations. For problems with nonlinear implicit model equations, the probably most effective approach is through parameter manipulation, such as shown in the previous example. A numerical example of another problem will be given in the next section.

6. EXAMPLE

We present as a numerical example a least squares model fitting problem arising from the adjustment of a planimetric net. The measurements in such a net have to satisfy net closure conditions, i.e., constraints without any parameters. The constraints form a set of simultaneous nonlinear equations involving all corrections, and net adjustment problems are not partitionable in this formulation. However, by introducing station coordinates as parameters net adjustment problems always can be partitioned.

A simple specific example is the planimetric traverse shown in Figure 1. Let the observations be the distances r_i between the stations and the corresponding azimuths ϕ_i. The constraints for the closed polygon of Figure 1 are obtained from the model equations (closure conditions)

$$\sum_{i=1}^{5} r_i \sin\phi_i = 0$$

and

$$\sum_{i=1}^{5} r_i \cos\phi_i = 0$$

(6.1)

by substituting in them the corrected $r_i + c_{ri}$ and $\phi_i + c_{\phi i}$ for the observed r_i and ϕ_i, respectively. Hence, the problem has two nonlinear scalar model equations (r=2), ten observations (n=10), and no parameters (p=0). For simplicity, we assume that the observations are not correlated, i.e., that the estimated variance-covariance matrix R is diagonal. Nevertheless, the adjustment problem is not partitionable in this formulation.

Next, we introduce, as parameters, the coordinates of the stations 1, 2, 3, and 4 relative to the reference station A. With these eight parameters, the two original model Eqs. (6.1) can be expressed equivalently by the following five sets of two equations each:

19

$$f_1 = \begin{pmatrix} r_1 \sin\phi_1 - x_1 \\ r_1 \cos\phi_1 - y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} ,$$

$$f_2 = \begin{pmatrix} x_1 + r_2 \sin\phi_2 - x_2 \\ y_1 + r_2 \cos\phi_2 - y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} ,$$

$$f_3 = \begin{pmatrix} x_2 + r_3 \sin\phi_3 - x_3 \\ y_2 + r_3 \cos\phi_3 - y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} , \qquad (6.2)$$

$$f_4 = \begin{pmatrix} x_3 + r_4 \sin\phi_4 - x_4 \\ y_3 + r_4 \cos\phi_4 - y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} ,$$

$$f_5 = \begin{pmatrix} x_4 + r_5 \sin\phi_5 \\ y_4 + r_5 \cos\phi_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

The corresponding constraint equations are obtained by substituting in Eq. (6.2) the corrected distances and angles for the observed ones. Now, we have s=5, r=10, n=10, and p=8. According to Eq. (5.2) no additional parameters can be introduced. In this formulation, the Jacobian matrix of the model function has the block diagonal form

$$\frac{\partial F}{\partial X} = \begin{pmatrix} \dfrac{\partial f_1}{\partial (r_1, \phi_1)} & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \dfrac{\partial f_5}{\partial (r_5, \phi_5)} \end{pmatrix} \qquad (6.3)$$

with 2 x 2 matrices in the diagonal. Hence, the problem partitions into five subsets.

20

We note in passing that in this example the reformulation of the model equations has only slightly reduced the size of the problem. In the original formulation, we would have to deal with 10 x 10 matrices in the normal equations, corresponding to the ten observations. The partitioning has reduced this part of the formulation to five 2 x 2 matrices, which can be handled much easier. However, we have also introduced eight parameters and corresponding 8 x 8 matrices in the normal equations. Hence, the reduction of matrix sizes is only from 10 x 10 to 8 x 8. If numerical efficiency were important in this example, then one would introduce fewer than eight parameters, e.g., only the coordinates of Stations 1 and 3. In that formulation, the largest matrix to be handled would be only 4 x 4.

Numerical values of the observations are given in Table 1, together with the results of the adjustments which were calculated with the computer program COLSMU of Reference 10. The parameter values, i.e., the station coordinates are listed in Table 2, and Table 3 provides the correlation coefficients between the station coordinates. The correlation coefficient matrix C is defined in terms of the variance-covariance matrix V_t by

$$C = D^{-1/2} V_t D^{-1/2} \tag{6.4}$$

where

$$D = \text{diag } V_t \; . \tag{6.5}$$

TABLE 1. OBSERVATIONS AND ADJUSTMENTS

Nr	r(km)	e_r	c_r	$r+c_r$	$\phi(°)$	e_ϕ	c_ϕ	$\phi+c_\phi$
1	10.5	0.47	0.356	10.856	77.0	1.0	0.33	76.67
2	2.9	0.25	-0.148	2.752	202.0	1.0	-0.01	201.99
3	4.6	0.32	-0.107	4.493	273.0	1.0	0.17	273.17
4	7.0	0.40	-0.199	6.801	151.0	1.0	-0.24	150.76
5	10.1	0.46	0.045	10.145	304.0	1.0	0.42	304.42

TABLE 2.   ADJUSTED COORDINATES OF STATIONS WITH ESTIMATED STANDARD ERRORS

Nr	x (km)	e_x(km)	y (km)	e_y(km)
1	10.563	0.324	2.503	0.183
2	9.533	0.306	−0.049	0.228
3	5.047	0.286	0.199	0.237
4	8.369	0.310	−5.735	0.242

Weighted sum of correction squares W = 1.665 635 46

Standard error with weight one m_o = 0.912 589

The factor m_o is not included in the standard errors of station coordinates.

TABLE 3.   CORRELATION MATRIX OF ADJUSTED STATION COORDINATES

	x_1	y_1	x_2	y_2	x_4	y_3	x_4	y_4
x_1	1.0000	0.2080	0.9572	−0.1477	0.5516	−0.1303	0.4494	−0.1801
y_1	0.2080	1.0000	0.1332	0.5161	−0.0594	0.4872	0.0394	0.1454
x_2	0.9572	0.1332	1.0000	0.0053	0.5358	0.0141	0.4626	−0.1375
y_2	−0.1477	0.5161	0.0053	1.0000	−0.2509	0.9417	−0.0407	0.3170
x_3	0.5516	−0.0594	0.5358	−0.2509	1.0000	−0.3059	0.8018	−0.3450
y_3	−0.1303	0.4872	0.0141	0.9417	−0.3059	1.0000	−0.0771	0.3470
x_4	0.4494	0.0394	0.4626	−0.0407	0.8018	−0.0771	1.000	−0.6092
y_4	−0.1801	0.1454	−0.1375	0.3170	−0.3450	0.3470	−0.6092	1.0000

Both matrices, C and V_t, were calculated by the cited utility routine which uses Eq. (3.5) for the calculation of V_t.

The adjusted positions of the stations are shown in Figure 1 together with corresponding two standard error ellipses indicating the accuracies of the positions.

Estimates of variances and covariances of the adjusted survey stations are important when the stations are used as bases for other measurements. As an example, let us assume that Stations 2 and 3 are used to determine the position of a target by azimuth measurements. Let ψ_2 and ψ_2 be the azimuths observed from Stations 2 and 3, respectively. Then the target is given by

$$
\left.
\begin{aligned}
x_t &= x_2 - \frac{\sin\psi_2}{\sin(\psi_3 - \psi_2)} \left[(x_3 - x_2)\cos\psi_3 + (y_3 - y_2)\sin\psi_3 \right] \\
y_t &= y_2 - \frac{\cos\sin\psi_2}{\sin(\psi_3 - \psi_2)} \left[(x_3 - x_2)\cos\psi_3 + (y_3 - Y_2)\sin\psi_3 \right]
\end{aligned}
\right\} \quad (6.6)
$$

Estimated variances and covariances of the target coordinates can be computed from the estimated accuracies of the azimuth observations and the variances and covariances of the bases' coordinates by applying the linearized law of variance propagation to Eq. (6.6). The results are shown in Table 4 and Figure 1. As expected, one obtains different estimates of the target accuracies, depending whether the correlations between stations are taken into account or not.

23

TABLE 4. TARGET POSITION

OBSERVED AZIMUTHS OF TARGET

$\psi_2 = 149.0 \pm 0.3$

$\psi_3 = 123.0 \pm 0.3$

COMPUTED COORDINATES OF TARGET

    a. All Covariances Considered

        $x_t = 12.159 \pm 0.437$

        $y_t = -4.419 \pm 0.337$

        Correlation coefficient $c_{xy} = -0.667256$

    b. Covariances Between Stations Neglected

        $x_t = 12.159 \pm 0.609$

        $y_t = -4.419 \pm 0.555$

        Correlation coefficient $c_{xy} = -0.906523$

    c. All Covariances Neglected

        $x_t = 12.159 \pm 0.629$

        $y_t = -4.419 \pm 0.616$

        Correlation coefficient $c_{xy} = -0.900271$
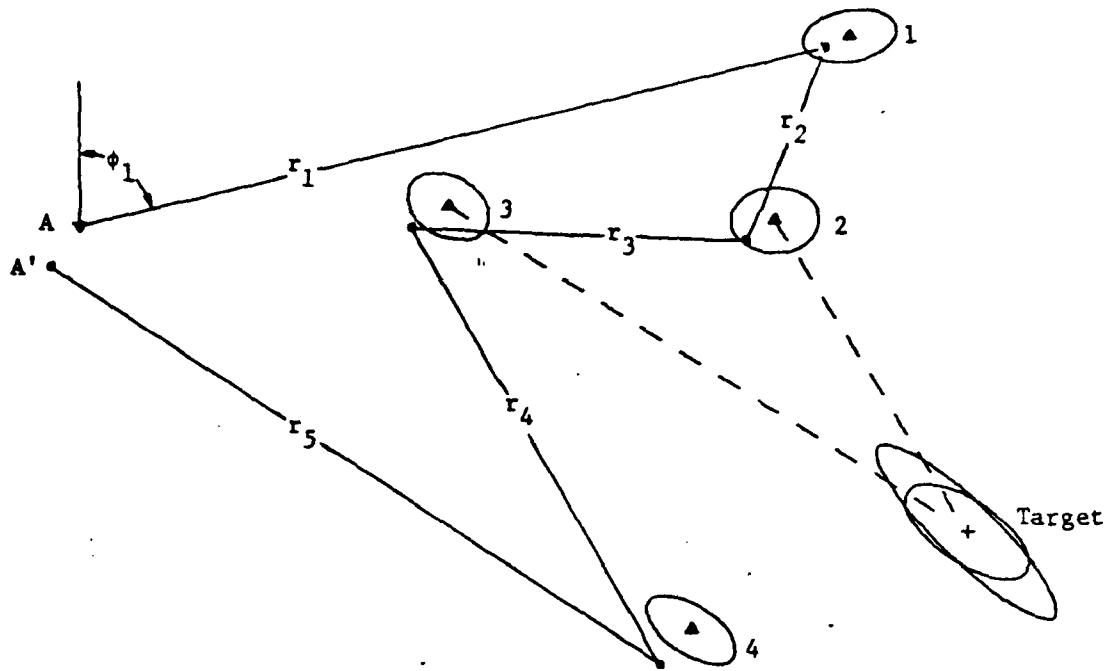
Figure 1. Planimetric Traverse

▼ – Reference station

■ – original survey station

▲ – adjusted survey station

+ – target, determined by azimuth measurements from Stations 2 and 3

The difference between Stations A and A' is the closure error of the traverse. A and A' coincide after adjustment. Accuracies of the adjusted stations and of the target are indicated by two standard error ellipses. The larger error ellipse around the target is obtained if correlations between station coordinates are neglected. (Case (c) in Table 4.) Data are given in Tables 1 through 4.

APPENDIX A

ITERATION FORMULAS

We provide a set of iteration formulas that are derived from the Newton Eqs. (3.1) by algebraic manipulations. First, we define the following matrices:

$$G = (F_X R F_X^T)^{-1} \qquad (A.1)$$

$$A = R F_X^T G F_X - I \qquad (A.2)$$

$$\Gamma = [I + AR(K^T F)_{XX}]^{-1} \qquad (A.3)$$

$$E_o = \Gamma \cdot [AC - R F_X^T G F_X] \qquad (A.4)$$

$$E_1 = \Gamma \cdot [R F_X^T G F_t + AR(K^T F)_{Xt}] \qquad (A.5)$$

$$D_o = (K^T F)_{tX} - F_t^T G F_X R(K^T F)_{XX} \qquad (A.6)$$

$$D_1 = (K^T F)_{tt} - F_t^T G F_X R(K^T F)_{Xt} \qquad (A.7)$$

$$N = F_t^T G F_t - D_1 + D_0 E_1 \qquad (A.8)$$

The iteration equations are

$$N\tau = F_t^T G(F_X C - F) + D_0 E_0 \qquad (A.9)$$

$$K + \kappa = G(F_X C - F) + G[F_t + F_X R(K^T F)_{Xt}] \tau - G F_X R(K^T F)_{XX} \epsilon \qquad (A.10)$$

$$\epsilon = E_0 - E_1 \tau . \qquad (A.11)$$

28

Numerical experiments have shown that the convergence of the iteration is enhanced if the equations are used in a subiteration mode by iterating alternatively on the parameters and residuals, respectively. For parameter subiteration only Eqs. (A.9) and (A.10) are used, assuming $\varepsilon \equiv 0$. For residual subiteration one sets $\tau \equiv 0$ and uses Eqs. (A.10) and (A.11).

In the variance formula (3.5) one uses N, defined by Eq. (A.8) and

$$S = F_t^T G F_X + D_0 \Gamma A \quad . \tag{A.12}$$

Another equivalent set of Newton-Raphson iteration equations is given in Reference 7. None of the sets is numerically superior to the other, and both require subiterations of parameters and residuals for efficiency.

Gauss-Newton iteration equations can be obtained from Newton-Raphson iteration equations by setting all second order derivatives equal to zero. The convergence of Gauss-Newton algorithms is inferior, but in some applications they have a larger domain of convergence.

LIST OF SYMBOLS

c, C - residuals

f, F - constraint (model) functions

k, K - Correlates (Lagrange multipliers)

M - norm matrix

n - dim X (total number of observations)

N - coefficient matrix for variance estimation of t

p - dim t (number of parameters)

r - dim F (number of scalar constraint equations)

R - estimated variance-covariance matrix of the observations X

s - number of subsets in a partitionable problem

S - influence matrix for variance estimation of t

t, T - parameters

x, X - observables

V_t - estimated variance-covariance matrix of t

W - object function $= c^T R^{-1} c$

$\hat{W}$ - modified object function

ε - correction of C

κ - correction of K

τ - correction of T

REFERENCES

1. W.E. Deming, "Statistical Adjustment of Data," John Wiley & Sons, New York, NY, 1944.

2. S. Brandt, "Statistical and Computational Methods in Data Analysis," North-Holland Publishing Co., Amsterdam, 1970.

3. D. Brown, "A Matrix Treatment of Least Squares Considering Correlated Observations," USA Ballistic Research Laboratories Report No. 937, 1955. (AD #71209)

4. J.M. Tienstra, "Theory of the Adjustment of Normally Distributed Observations," N.V. Uitgevery "Argus," Amsterdam, 1956.

5. M.F. Britt and R.H. Luecke, "The Estimation of Parameters in Non-linear, Implicit Models," Technometrics, 15, 233-247, 1973.

6. A. Celmiņš, "Least Squares Adjustment with Finite Residuals for Non-Linear Constraints and Partially Correlated Data," USA Ballistic Research Laboratories Report No. R-1658, 1973. (AD #766283)

7. A.F. Pope, "Two Approaches to Nonlinear Least Squares Adjustments," The Canadian Surveyor, 28, 663-669, 1974.

8. R.M. Passi, "Use of Nonlinear Least Squares in Meterological Applications," Journal of Applied Meterology, 16, 828-832, 1977, and 17, 1579-1580, 1978.

9. W.H. Jefferys, "On the Method of Least Squares," The Astronomical Journal, 85, 177-182, 1980.

10. A. Celmiņš, "A Manual for General Least Squares Model Fitting," USA Ballistic Research Laboratory Report ARBRL-TR-02167, 1979. (AD #B040229L)

DISTRIBUTION LIST

No. of Copies	Organization	No. of Copies	Organization
12	Commander Defense Technical Info Center ATTN: DDC-DDA Cameron Station Alexandria, VA 22314	1	Commander US Army Communications Research and Development Command ATTN: DRDCO-PPA-SA Fort Monmouth, NJ 07703
1	Commander US Army Materiel Development and Readiness Command ATTN: DRCDMD-ST 5001 Eisenhower Avenue Alexandria, VA 22333	1	Commander US Army Electronics Research and Development Command Technical Support Activity ATTN: DELSD-L Fort Monmouth, NJ 07703
2	Commander US Army Armament Research and Development Command ATTN: DRDAR-TSS Dover, NJ 07801	2	Commander US Army Missile Command ATTN: DRSMI-R DRSMI-YDL Redstone Arsenal, AL 35809
1	Commander US Army Armament Materiel Readiness Command ATTN: DRSAR-LEP-L, Tech Lib Rock Island, IL 61299	2	Commander US Army Tank Automotive Research and Development Command ATTN: DRDTA-UL DRDTA-ZSS, T. Szako Warren, MI 48090
1	Director US Army ARRADCOM Benet Weapons Laboratory ATTN: DRDAR-LCB-TL Watervliet, NY 12189	1	Director US Army TRADOC Systems Analysis Activity ATTN: ATAA-SL, Tech Lib White Sands Missile Range NM 88002
1	Commander US Army Aviation Research and Development Command ATTN: DRSAV-E P.O. Box 209 St. Louis, MO 61366	1	Commander Naval Research Laboratory Marine Technology Division ATTN: Structural Reliability Section, J. Sanford Washington, DC 20375
1	Director US Army Air Mobility Research and Development Laboratory Ames Research Center Moffett Field, CA 94035		

DISTRIBUTION LIST

No. of Copies	Organization	No. of Copies	Organization
3	National Bureau of Standards ATTN: W.S. Horton D. Kahaner D. Shier Washington, DC 20234	1	Union Carbide Corporation Chemicals and Plastics ATTN: H.J. Britt P.O. Box 8361 Charleston, WV 25303
1	National Center of Atmospheric Research ATTN: R.M. Passi Boulder, CO 80803	1	Cornell University Computer Science Department ATTN: J. Dennis Ithaca, NY 14850
1	National Geodetic Survey Geodetic Research and Development Laboratory ATTN: A.J. Pope Rockville, MD 20852	1	Florida State University Institute of Molecular Biophysics Department of Chemistry ATTN: D.F. DeTar Tallahassee, FL 32306
1	US Geological Survey ATTN: J.R. Fisher Washington, DC 20242	1	The Johns Hopkins University Mathematical Sciences ATTN: R.H. Byrd Baltimore, MD 21218
1	Brookhaven National Laboratory Associated Universities, Inc. ATTN: W.H. Sachs Upton, Long Island, NY 11973	1	Pennsylvania State University Computer Sciences Department ATTN: P.A.D. DeMaine University Park, PA 16807
1	Lawrence Livermore National Laboratory University of California ATTN: R.M. Singleton P.O. Box 5508 Livermore, CA 94550	1	Ohio State University Department of Geodetic Science ATTN: R.H. Rapp 1958 Neil Avenue Columbus, OH 43210
2	The Perkin Elmer Corporation ATTN: E. Niple R.J. Noll Main Avenue Norwalk, CT C6986	1	Texas A&M University ATTN: S.B. Childs College Station, TX 77843
1	SKF Industries ATTN: J.I. McCool 1100 First Avenue King of Prussia, PA 19406	1	State University of New York Department of Earth and Space Sciences ATTN: D.M. Peterson, Stony Brook, NY 12222

DISTRIBUTION LIST

No. of Copies	Organization	No. of Copies	Organization

1	University of California
	Department of Chemistry
	ATTN: J.R. Murdoch
	Los Angeles, CA 90024
			Aberdeen Proving Ground
			  Dir, USAMSAA
			    ATTN: DRXSY-D
			           DRXSY-MP, H. Cohen
			  Cdr, USATECOM
			    ATTN: DRSTE-TO-F
			  Dir, USACSL, Bldg. E3516, EA
			    ATTN: DRDAR-CLB-PA

1	University of California
	School of Veterinary Science
	Radiobiology Laboratory and
	  Department of Radiological
	  Sciences
	ATTN: O.G. Raabe
	Davis, CA 95616		

1	University of Colorado Computer Science Department ATTN: R. Schnabel Boulder, CO 80304		
1	University of Maryland Baltimore County Department of Mathematics ATTN: P.R. Kumar 5401 Wilkins Avenue Baltimore, MD 21228		
1	University of Texas Department of Astronomy ATTN: W.H. Jefferys Austin, TX 78712		
1	National Heart and Lung Institute ATTN: Biometrics Rsch Branch, J.H. Ware Bethesda, MD 20014		
1	National Institute of Health ATTN: R.I. Shrager 900 Rockville Pike Bethesda, MD 20205		

USER EVALUATION OF REPORT

Please take a few minutes to answer the questions below; tear out
this sheet, fold as indicated, staple or tape closed, and place
in the mail. Your comments will provide us with information for
improving future reports.

1. BRL Report Number_____

2. Does this report satisfy a need? (Comment on purpose, related
project, or other area of interest for which report will be used.)

_____

_____

_____

3. How, specifically, is the report being used? (Information
source, design data or procedure, management procedure, source of
ideas, etc.)_____

_____

_____

4. Has the information in this report led to any quantitative
savings as far as man-hours/contract dollars saved, operating costs
avoided, efficiencies achieved, etc.? If so, please elaborate.

_____

_____

5. General Comments (Indicate what you think should be changed to
make this report and future reports of this type more responsive
to your needs, more usable, improve readability, etc.)_____

_____

_____

_____

6. If you would like to be contacted by the personnel who prepared
this report to raise specific questions or discuss the topic,
please fill in the following information.
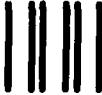
                    Name:_____

          Telephone Number:_____

Organization Address:_____

                      _____

                      _____